TIGHT IS BETTER: PERFORMANCE IMPROVEMENT OF THE COMPRESSIVE CLASSIFIER USING EQUI-NORM TIGHT FRAMES

Hailong Shi and Hao Zhang

Department of Electronic Engineering, Tsinghua University, Beijing

ABSTRACT

Detecting or classifying already known sparse signals contaminated by Gaussian noise from compressive measurements is different from reconstructing sparse signals, as its objective is to minimize the error probability which describes performance of the detectors or classifiers. This paper is concerned about the performance improvement of a commonly used Compressive Classifier. We prove that when the arbitrary sensing matrices used to get the Compressive Measurements are transformed into Equi-Norm Tight Frames, i.e. the matrices that are row-orthogonal, The Compressive Classifier achieves better performance. Although there are other proofs that among all Equi-Norm Tight Frames the Equiangular tight Frames (ETFs) bring best worst-case performance, the existence and construction of ETFs on some dimensions is still an open problem. As the construction of Equi-Norm Tight Frames from any arbitrary matrices is very easy and practical compared with ETF matrices, the result of this paper can also provide a practical method to design an improved sensing matrix for Compressive Classification. We can conclude that: Tight is Better!

***Index Terms*—** Compressive Classification, Sensing Matrix, Tight Frame, Equiangular Tight Frame (ETF)

1. INTRODUCTION

Recent years considerable progress has been made towards sensing and recovery of sparse signals, known as Compressive Sensing [1][2]. However, less attention has been paid to detecting or classifying sparse signals from compressive measurements, which can be stated as the problem of Compressive Detection or Compressive Classification.

The problem of Compressive Detection and Compressive Classification is derived from the theory of statistical signal processing [3]. In this compressive scenario, Gaussian contaminated sparse signals are compressed by under-determined linear transformations, and then the compressed measurements are used for signal detection or classification. Since the objective in Compressive Detection or Classification is to minimize the error probability of the detectors or classifiers, the analysis of the influence of sensing matrices on the Compressive Detection or Classification is different from that of reconstructing sparse signals.

The earliest analysis of the performance of Compressive Detection and Compressive Classification we ever know is from Davenport's publications in 2006 [4][5]. In these publications the row-orthogonal random projection matrices were analyzed to get some performance bounds on the error probability of Compressive Detectors and Classifiers.

Besides, there are other publications such as Ramin Zahedi *et al.* 's works [6][7], in which Equiangular Tight Frames (ETFs, [8]) were considered and proved to have best worst-case performance in Compressive Detection and Classification under the constraint that the sensing matrices are row-orthogonal.

This paper is mainly concerned about the design of sensing matrices to improve the performance of Compressive Classification. We noticed that all these publications mentioned above ([4][5][6][7]) and many other related most recent researches ([9][10]) did their analysis under the constraint that the sensing matrices are row-orthogonal, *i.e.* the column vectors form an Equi-Norm Tight Frame. The reason why they use this constraint is for simplicity and avoiding the coloring of received signals between different hypotheses. While in this paper, we will show that the transformation to Equi-Norm Tight Frames from arbitrary sensing matrices will make the error probability of the commonly used Compressive Classifier to decrease, which coincides with the row-orthogonal constraint commonly used above in ([4][5][6][7]). Although there are already proofs ([6][7]) showing that Equiangular Tight Frames (ETFs) have best worst-case performance among all Tight Frames, our job is different from theirs. Because constructing Equi-Norm Tight Frames is much easier and more practical, our conclusion can be: Tight is Better at both improved performance and practical construction!

The remainder of this paper is organized as follows: In Section 2, we give the problem formulation of Compressive Classification and the error probability; in Section 3, we demonstrate our main result and proof of the influence of Equi-Norm Tight Frames on the error probability; and in the last section simulation results and discussions are provided.

This work was supported in part by the National Basic Research Program of China (973 Program, No. 2010CB731901) and in part by the National Natural Science Foundation of China (No. 40901157 and No. 61201356).

2. COMPRESSIVE SIGNAL CLASSIFICATION

Derived from the traditional theory of statistical signal detection and estimation ([3]), in the Compressive Classification scenario, the sparse signals is compressed by under-determined linear transformation first, then these compressed measurements are used to do classification. As is stated in Davenport's works ([4][5]), the model of compressive signal classification can be expressed as:

$$\boldsymbol{y} = \begin{cases} \boldsymbol{\Phi}(\boldsymbol{s_1} + \boldsymbol{n}) & \text{Hypothesis } H_1 \\ \boldsymbol{\Phi}(\boldsymbol{s_2} + \boldsymbol{n}) & \text{Hypothesis } H_2 \\ \cdots & \cdots \\ \boldsymbol{\Phi}(\boldsymbol{s_m} + \boldsymbol{n}) & \text{Hypothesis } H_m \end{cases} \quad (1)$$

in which $\boldsymbol{s_i} \in \Lambda_k = \{\boldsymbol{s} \in \mathbb{R}^N, \|\boldsymbol{s}\|_0 \le k\}$ are k-sparse signals, i.e. the vectors with at most k nonzero elements, besides they are orthogonal to each other and have equal norms ($\langle \boldsymbol{s_i}, \boldsymbol{s_j}\rangle = 0, \|\boldsymbol{s}_i\|_2^2 = \|\boldsymbol{s}_j\|_2^2, i \neq j$). $\boldsymbol{y} \in \mathbb{R}^n$ is the received compressive measurements, $\boldsymbol{n} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I_N})$ is the Gaussian noise contaminating the sparse signals. And $\boldsymbol{\Phi} \in \mathbb{R}^{n\times N}, n < N$ is the under-determined sensing matrix, which has full row rank and equal column norms, and also satisfies RIP property ([11][12]). All those conditions above are generally used for convenience or necessity.

According to the works of Davenport [5], Zahedi [6] and Rao [9], under the constraint of row-orthogonal ($\boldsymbol{\Phi\Phi^T = I}$) to the sensing matrices, the following classifier can be used:

$$\begin{aligned} t_i &= \boldsymbol{y^T \Phi s_i} - \frac{1}{2}\|\boldsymbol{\Phi s_i}\|_2^2 \\ &= \langle \boldsymbol{y}, \boldsymbol{\Phi s_i}\rangle - \frac{1}{2}\langle \boldsymbol{\Phi s_i}, \boldsymbol{\Phi s_i}\rangle, \quad i = 1, 2, \cdots, m \end{aligned} \quad (2)$$

And the classification result is

$$i^* = \arg\max_i \{t_i\} \quad i = 1, 2, \cdots, m \quad (3)$$

The classifier in (2) is a most commonly used classifier, it is composed of a term of correlation and an minus term for normalization. The correlation part $\boldsymbol{y^T \Phi s_i}$ also plays an important role in iterations of the Orthogonal Matching Pursuit (OMP) algorithm [13]. Since all the known analysis of Compressive Classifiers is restricted to row-orthogonal sensing matrices, our work is going to break this restriction. So the performance of the Compressive Classifier in (2) with more general sensing matrices that may or may not be row-orthogonal will be analyzed as follows.

Theorem 1. *If we use Compressive Classifier (2) and (3) to classify (1), the error probability can be expressed as:*

$$\mathcal{P}_E \le \sum_{i \neq T}^{m} \Phi\Big(-\frac{\|\boldsymbol{\Phi(s_T - s_i)}\|_2^2}{2\sigma\|\boldsymbol{\Phi^T\Phi(s_T - s_i)}\|_2}\Big) \quad 1 \le T \le m \quad (4)$$

Here $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp{(-\frac{t^2}{2})}\mathrm{d}t$*, and* $\boldsymbol{\Phi} \in \mathbb{R}^{n\times N}, n < N$ *is the sensing matrix.*

Proof. The error probability is [3]

$$\mathcal{P}_E = 1 - \sum_{i=1}^{m} \mathcal{P}(t_i > \max_{1\le j\le m, j\neq i}\{t_j\}|H_i)\mathcal{P}(H_i) \quad (5)$$

In general, the prior probability of different hypotheses can be assumed as uniformly distributed ($\mathcal{P}(H_i) = \mathcal{P}(H_j), i \neq j$), and the conditional probability in (5) is assumed to be equal [3] for symmetry, that is

$$\begin{aligned} &\mathcal{P}(t_i > \max_{1\le k\le m, k\neq i}\{t_k\}|H_i) = \\ &\mathcal{P}(t_j > \max_{1\le k\le m, k\neq j}\{t_k\}|H_j) \qquad i \neq j \end{aligned} \quad (6)$$

then the error probability is

$$\begin{aligned} \mathcal{P}_E &= 1 - \mathcal{P}(t_T > \max_{1\le i\le m, i\neq T}\{t_i\}|H_T) \\ &= 1 - \mathcal{P}(t_T > t_i, \forall i \neq T | H_T) \end{aligned} \quad (7)$$

Thus the error probability can be analyzed with respect to the conditional probability $\mathcal{P}(t_T > t_i, \forall i \neq T|H_T)$ under the H_T hypothesis. The statistics (2) under the H_T hypothesis have the following joint distributions

$$[t_T \quad t_i]^T \sim \mathcal{N}(\boldsymbol{\mu}_{T,i}, \boldsymbol{\Sigma}_{T,i}) \quad (8)$$

and

$$\boldsymbol{\mu}_{T,i} = \begin{bmatrix} \frac{1}{2}\|\boldsymbol{\Phi s_T}\|_2^2 \\ \langle \boldsymbol{\Phi s_i}, \boldsymbol{\Phi s_T}\rangle - \frac{1}{2}\|\boldsymbol{\Phi s_i}\|_2^2 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_{T,i} = \sigma^2 \begin{bmatrix} \|\boldsymbol{\Phi^T\Phi s_T}\|_2^2 & \langle \boldsymbol{\Phi^T\Phi s_T}, \boldsymbol{\Phi^T\Phi s_i}\rangle \\ \langle \boldsymbol{\Phi^T\Phi s_i}, \boldsymbol{\Phi^T\Phi s_T}\rangle & \|\boldsymbol{\Phi^T\Phi s_i}\|_2^2 \end{bmatrix}$$

Combined with the Union Bound of probability theory, the error probability is then

$$\mathcal{P}_E \le \sum_{i \neq T}^{m} \Phi\Big(-\frac{\|\boldsymbol{\Phi(s_T - s_i)}\|_2^2}{2\sigma\|\boldsymbol{\Phi^T\Phi(s_T - s_i)}\|_2}\Big) \quad 1 \le T \le m$$

and $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp{(-\frac{t^2}{2})}\mathrm{d}t$. □

As a matter of fact, if there is only two hypotheses in the Compressive Classification problem (1), the inequality becomes equality in (4).

So in order to decrease the error probability (4) of classifier (2) without the constraint of row-orthogonal to sensing matrices and for all possible k-sparse signals $\boldsymbol{s}_i$, we will have to increase

$$\frac{\|\boldsymbol{\Phi(s_T - s_i)}\|_2^2}{\|\boldsymbol{\Phi^T\Phi(s_T - s_i)}\|_2}, \quad 1 \le T \neq i \le m \quad (9)$$

for all k-sparse signals $\boldsymbol{s_T}, \boldsymbol{s_i} \in \Lambda_k = \{\boldsymbol{s} \in \mathbb{R}^N, \|\boldsymbol{s}\|_0 \le k\}$.

In the cases where sensing matrices satisfying row-orthogonal ($\boldsymbol{\Phi\Phi^T} = I$), (9) is then reduced to

$$\frac{\|\boldsymbol{\Phi(s_T - s_i)}\|_2^2}{\|\boldsymbol{\Phi^T\Phi(s_T - s_i)}\|_2} = \|\boldsymbol{\Phi(s_T - s_i)}\|_2, \quad 1 \le T \neq i \le m \quad (10)$$

And this is what Davenport [4][5] and Zahedi [6][7] analyzed in their publications.

3. SENSING MATRICES AND THE ERROR PROBABILITY OF COMPRESSIVE SIGNAL CLASSIFICATION

Although there has been plenty of works about the performance analysis of Compressive Detection and Classification, all these works have the same row-orthogonal presumption, and without a theoretical explanation. However, what we believe is that there exist other important reasons for the row-orthogonal condition to be necessary. Here is our main result of this paper:

Theorem 2. *In the Compressive Classification problem (1), If the sensing matrix $\mathbf{\Phi} \in \mathbb{R}^{n\times N}, n < N$ is transformed into an Equi-Norm Tight Frame, the error probability (4) of the classifier (2) will decrease.*

Proof. According to Section 2, the error probability (4) is determined by the following expression (9):

$$\frac{\|\mathbf{\Phi}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\mathbf{\Phi^T}\mathbf{\Phi}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2} \quad 1 \le T \ne i \le m$$

for all k-sparse signals $\boldsymbol{s}_T, \boldsymbol{s}_i$, where $\mathbf{\Phi}$ is an arbitrary sensing matrix satisfying RIP.

According to the basic presumptions of $\mathbf{\Phi}$ in (1), the arbitrary under-determined sensing matrix $\mathbf{\Phi} \in \mathbb{R}^{n\times N}$, $n < N$ has full row rank, thus the singular value decomposition of $\mathbf{\Phi}$ is

$$\mathbf{\Phi} = \boldsymbol{U}\left[\mathbf{\Sigma_n} \quad \boldsymbol{O}\right]\boldsymbol{V^T} \tag{11}$$

Here $\mathbf{\Sigma}_n \in \mathbb{R}^{n\times n}$ is a diagonal matrix with each element $\mathbf{\Phi}$'s singular value $\sigma_i \ne 0$ $(1 \le i \le n)$, and $\boldsymbol{U} \in \mathbb{R}^{n\times n}$, $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ are orthogonal matrices composed of $\mathbf{\Phi}$'s left and right singular vectors.

If an arbitrary equi-norm sensing matrix $\mathbf{\Phi}$ is transformed into an Equi-Norm Tight Frame $\hat{\mathbf{\Phi}}$, we do row-orthogonalization to its row vectors, which is equivalent as:

$$\hat{\mathbf{\Phi}} = \sqrt{c}\cdot\boldsymbol{U}\mathbf{\Sigma_n^{-1}}\boldsymbol{U^T}\mathbf{\Phi} = \sqrt{c}\cdot\boldsymbol{U}\left[\boldsymbol{I_n} \quad \boldsymbol{O}\right]\boldsymbol{V^T} \tag{12}$$

Where $\boldsymbol{U} \in \mathbb{R}^{n\times n}$, $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ are $\mathbf{\Phi}$'s singular vector matrices. In a word, row-orthogonalization is equivalent to transforming all singular values of $\mathbf{\Phi}$ into equal ones. Thus $\hat{\mathbf{\Phi}}\hat{\mathbf{\Phi}}^{\boldsymbol{T}} = \boldsymbol{c}\cdot\boldsymbol{I_n}$, where $c > 0$ is a certain constant for normalization.

Then

$$\frac{\|\hat{\mathbf{\Phi}}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\hat{\mathbf{\Phi}}^{\boldsymbol{T}}\hat{\mathbf{\Phi}}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2} = \|\left[\begin{array}{cc}\boldsymbol{I_n} & \boldsymbol{O}\end{array}\right]\boldsymbol{V}^T(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2 \tag{13}$$

And for arbitrary sensing matrix $\mathbf{\Phi}$ that may not be row-orthogonal, we have

$$\begin{aligned}&\frac{\|\mathbf{\Phi}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\mathbf{\Phi^T}\mathbf{\Phi}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2} \\ &= \frac{\|\left[\begin{array}{cc}\mathbf{\Sigma}_n & \boldsymbol{O}\end{array}\right]\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\left[\begin{array}{cc}\mathbf{\Sigma}_n^2 & \boldsymbol{O}\end{array}\right]\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2}\end{aligned} \tag{14}$$

If we denote $\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i})$ by $\boldsymbol{u}^{(T,i)}$, where $\boldsymbol{u}^{(T,i)} = [u_1, u_2, \cdots, u_N]^T$. Then (14) becomes

$$\begin{aligned}\frac{\|\left[\begin{array}{cc}\mathbf{\Sigma}_n & \boldsymbol{O}\end{array}\right]\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\left[\begin{array}{cc}\mathbf{\Sigma}_n^2 & \boldsymbol{O}\end{array}\right]\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2} &= \frac{\sum_{j=1}^n \sigma_j^2 u_j^2}{\sqrt{\sum_{j=1}^n \sigma_j^4 u_j^2}} \\ &\le \sqrt{\sum_{j=1}^n u_j^2} = \|\left[\begin{array}{cc}\boldsymbol{I_n} & \boldsymbol{O}\end{array}\right]\boldsymbol{V}^T(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2\end{aligned} \tag{15}$$

The last inequality is derived from the Cauchy-Schwarz Inequality, combining (14), (15) with (13), then we have

$$\frac{\|\mathbf{\Phi}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\mathbf{\Phi^T}\mathbf{\Phi}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2} \le \frac{\|\hat{\mathbf{\Phi}}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2^2}{\|\hat{\mathbf{\Phi}}^{\boldsymbol{T}}\hat{\mathbf{\Phi}}(\boldsymbol{s_T}-\boldsymbol{s_i})\|_2} \tag{16}$$

which means that row-orthogonalization makes (9) larger and thus brings lower error probability. The condition when the equality holds is that

$$\left[\mathbf{\Sigma_n^2} \quad \boldsymbol{O}\right]\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i}) = c\cdot\left[\boldsymbol{I_n} \quad \boldsymbol{O}\right]\boldsymbol{V^T}(\boldsymbol{s_T}-\boldsymbol{s_i}) \tag{17}$$

where $c > 0$ is a certain constant.

It is obvious that the equality in (16) holds for all k-sparse signals $\boldsymbol{s}_T$ and $\boldsymbol{s}_i$, if and only if $\mathbf{\Sigma_n^2} = \boldsymbol{c}\cdot\boldsymbol{I_n}$, which means

$$\mathbf{\Phi}\mathbf{\Phi^T} = c\cdot\boldsymbol{I}_n \tag{18}$$

So the result of (16) means that when arbitrary under-determined sensing matrices $\mathbf{\Phi} \in \mathbb{R}^{n\times N}, n < N$ are transformed into an Equi-Norm Tight Frame, i.e. row-orthogonalized, the equality in (16) will hold, then the value of (9) will get larger, which also means improvement of the performance of Compressive Classifier (2). □

the constant $c > 0$ above is an amplitude constant for normalization, with the equi-norm presumption of sensing matrices, the following corollary can be deduced:

Corollary 1. *If a matrix $\mathbf{\Phi} \in \mathbb{R}^{n\times N}, n < N$ form an Equi-Norm Tight Frame, that is $\mathbf{\Phi}\mathbf{\Phi^T} = c\cdot I_n$ and $\|\boldsymbol{\phi}_i\|_2 = \|\boldsymbol{\phi}_j\|_2 = \psi$, then $c = \frac{N}{n}\psi^2$.*

Proof. If $\mathbf{\Phi}$ has equal column norms and satisfies $\mathbf{\Phi}\mathbf{\Phi^T} = c\cdot I_n$, then

$$tr(\mathbf{\Phi^T}\mathbf{\Phi}) = N\cdot\psi^2 = tr(\mathbf{\Phi}\mathbf{\Phi^T}) = n\cdot c \tag{19}$$

As a result, $c = \frac{N}{n}\psi^2$. □

If we let $c = 1$, then we can get $\|\boldsymbol{\phi}_i\|_2 = \|\boldsymbol{\phi}_j\|_2 = \sqrt{n/N}$, which coincides with the results of [6] and [7].

Before the end of this section, some important discussions are believed to be necessary here.

- Remark 1. The result of Theorem 2 and the inequality (16) does not mean that Equi-Norm Tight Frames

as the sensing matrices will make the error probability of the commonly used Compressive Classifier (2) achieve an optimal lower bound. Actually the right side of (16) is still dependent on $\mathbf{\Phi}$ because of the existence of $\mathbf{\Phi}$'s right singular vector matrix $\boldsymbol{V}$. Our point in theorem 2 is that when arbitrary sensing matrices $\mathbf{\Phi} \in \mathbb{R}^{n\times N}, n < N$ are "tightened", or transformed into Equi-Norm Tight Frames, then the equality in (16) holds; as a result, the error probability (4) will become lower, thus brings performance improvement for the Compressive Classifier (2). Similar discussions about the improvement of Equi-Norm Tight Frames to oracle estimators can be found in [14], which is also a good support of the advantage of "Tight". In a word, Tight is better!

- Remark 2. When the inequality (16) becomes equality, that is when the sensing matrices satisfy $\mathbf{\Phi\Phi^T} = \boldsymbol{c \cdot I}$, it can be proved that the error probability satisfies

$$\mathcal{P}_E \leq \sum_{i\neq T}^{m} \Phi(-\frac{\|\mathbf{\Phi}(\boldsymbol{s_T} - \boldsymbol{s_i})\|_2}{2c^{-1/2}\cdot\sigma}) \quad 1 \leq T \leq m$$

  which coincides with Davenport's [4] and [5] and Zahedi's [6] and [7], where they set $c = 1$. So this also explains the benefits of using the row-orthogonal constraint to do Compressive Classification.

- Remark 3. In comparison with the the work of Zahedi in [6] and [7], where Equiangular Tight Frames (ETFs) are proved to have the best worst-case performance among all Tight Frames (row-orthogonal constrained matrices), we just give the proof that for general under-determined sensing matrices, tightening can bring performance improvement for Compressive Classification. Our job is different from theirs, because all of Zahedi's analysis is based on the constraint that the sensing matrices are tight, or row-orthogonal, and the advantage of Equiangular Tight Frames (ETFs, [8]) is that ETFs have the best worst-case (maximum of the minimum) performance among all Tight Frames of same dimensions, while our result shows that when arbitrary sensing matrices is "tightened", i.e. transformed into Equi-Norm Tight Frames, the performance of Compressive Classification will get improved. Nonetheless, the existence and construction of ETFs of some certain dimensions remains an open problem ([8]), while doing row-orthogonalization for any arbitrary matrix is very easy and practical. So our point is that Tight is Better at both improved performance and practical construction.

4. SIMULATIONS

In this section the main result of theorem 2 is verified by Monte-Carlo simulations. In the simulation some arbitrary

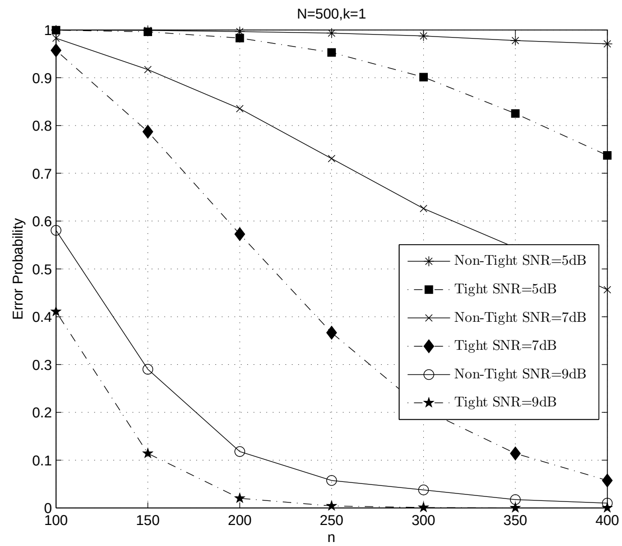

Fig. 1. Monte-Carlo simulation of classification error probability using non-tight frames and tight frames

$k = 1$ sparse signals are generated and classified using Non-Tight Frames and Tight Frames. Here we choose $N = 500$, and the error probabilities is plotted in Fig.1 corresponding with number of measurements n from 100 to 400 and signal to noise ratios $\|\boldsymbol{s_i}\|_2/\sigma$ from 5 dB to 9 dB. Each error probability is calculated from average of 5000 independent experiments with 10 instances of tight or non-tight sensing matrices.

The simulation shows that Tight Frames transformed from general Gaussian random matrices have better Compressive Classification performance within n's whole range, which is the benefits that "Tightening" brings.

5. CONCLUSION

This paper deals with the performance improvement of a commonly used Compressive Classifier (2). We prove that the transformation to Equi-Norm Tight Frames from arbitrary sensing matrices will make the error probability of the commonly used Compressive Classifier to decrease, thus improve the classification performance, which coincides with the row-orthogonal constraint commonly used before. Although there are other proofs that among all Equi-Norm Tight Frames the Equiangular tight Frames (ETFs) achieves best worst-case classification performance, the existence and construction of ETFs of some dimensions is still an open problem. As the construction of Equi-Norm Tight Frames from any arbitrary matrices is very easy and practical, the conclusion of this paper can also provide a practical method to design an improved sensing matrix for Compressive Classification. In a word, we can conclude that Tight is Better!

6. REFERENCES

[1] Emamnuel J. Candes, "Compressive sampling," 2006.

[2] E.J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489 – 509, feb. 2006.

[3] Steven M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*, Prentice Hall PTR, 1998.

[4] Mark A. Davenport, Michael B. Wakin, and Richard G. Baraniuk, "Detection and estimation with compressive measurements," Tech. Rep., 2006.

[5] M.A. Davenport, P.T. Boufounos, M.B. Wakin, and R.G. Baraniuk, "Signal processing with compressive measurements," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 445 –460, april 2010.

[6] R. Zahedi, A. Pezeshki, and E.K.P. Chong, "Robust measurement design for detecting sparse signals: Equiangular uniform tight frames and grassmannian packings," in *American Control Conference (ACC), 2010*, 30 2010-july 2 2010, pp. 4070 –4075.

[7] Ramin Zahedi, Ali Pezeshki, and Edwin K.P. Chong, "Measurement design for detecting sparse signals," *Physical Communication*, vol. 5, no. 2, pp. 64 – 75, 2012.

[8] Shayne Waldron, "On the construction of equiangular frames from graphs," *Linear Algebra and its Applications*, vol. 431, no. 11, pp. 2228 – 2242, 2009.

[9] B.S.M.R. Rao, S. Chatterjee, and B. Ottersten, "Detection of sparse random signals using compressive measurements," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, march 2012, pp. 3257 –3260.

[10] T. Wimalajeewa, Hao Chen, and P.K. Varshney, "Performance limits of compressive sensing-based signal classification," *Signal Processing, IEEE Transactions on*, vol. 60, no. 6, pp. 2758 –2770, june 2012.

[11] E. J. Candes and T Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215–, 2005.

[12] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, pp. 253–263, 2008, 10.1007/s00365-007-9003-x.

[13] M.A. Davenport and M.B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4395 –4401, sept. 2010.

[14] Wei Chen, M.R.D. Rodrigues, and I.J. Wassell, "On the use of unit-norm tight frames to improve the average mse performance in compressive sensing applications," *Signal Processing Letters, IEEE*, vol. 19, no. 1, pp. 8 –11, jan. 2012.